Three eras of digital governance

Jonathan Zittrain 2019-11-27T10:00:53

To understand where digital governance is going, we must take stock of where it's been, because the timbre of mainstream thinking around digital governance today is dramatically different than it was when study of "Internet governance" coalesced in the late 1990s. Perhaps the most obvious change has been from emphasizing networked technologies' positive effects and promise – couched around concepts like connectivity, innovation, and, by this author, "generativity" – to pointing out their harms and threats. It's not that threats weren't previously recognized, but rather that they were more often seen in external clamps on technological development and upon the corresponding new freedoms for users, whether government intervention to block VOIP services like Skype to protect incumbent telco revenues, or in the shaping of technology to effect undue surveillance, whether for government or corporate purposes. The shift in emphasis from positive to negative corresponds to a change in the overarching frameworks for talking about regulating information technology. We have moved from a discourse around *rights* – particularly those of end-users, and the ways in which abstention by intermediaries is important to facilitate citizen flourishing – to one of *public health*, which naturally asks for a weighing of the systemic benefits or harms of a technology, and to think about what systemic interventions might curtail its apparent excesses. Each framework captures important values around the use of technology that can both empower and limit individual freedom of action, including to engage in harmful conduct. Our goal today should be to identify where competing values frameworks themselves preclude understanding of others' positions about regulation, and to see if we can map a path forward that, if not reconciling the frameworks, allows for satisfying, if ever-evolving, resolutions to immediate questions of public and private governance.

The rights era

The original consideration of threats as external to the otherwise-mostly-beneficial uses of tech made for a ready framing of Internet governance issues around rights, and in particular a classic libertarian ethos of the preservation of rapidly-growing individual affordances in speech – "now anyone can speak without a gatekeeper!" – against encroachment by government censorship or corporate pushback motivated by the disruption of established business models.

A good example in the first category are the debates around the U.S. Communications Decency Act of 1995 (CDA) , which sought to keep indecent material away from minors by penalizing those who indiscriminately made it available online. The Supreme Court struck down the core provisions of the CDA in 1997 on First Amendment grounds, holding that too much protected speech would be chilled by the law, and successor laws met a similar fate. Another example can be found in the early and then not-officially-acknowledged efforts by the Chinese government to block citizens' access to web sites critical of the state, something viewed among

those studying Internet governance as an unalloyed wrong, not least because of the lack of due process, including notification, in effecting any blocks.

When the Internet's affordances for near-instant file transfer led to objections by publishers and other copyright holders over copyright infringement, those against stepped-up enforcement or new requirements for intermediaries relied on a rights-centric account. Copyright itself establishes legally protected interests – rights – but the sorts of interventions required to continue to secure those rights in practice were described early and often as overly burdening individual rights, whether through content takedown schemes to be effectuated by intermediaries, or individual lawsuits filed against those engaged in the sharing of copyright material.

It is in intermediary liability that the most significant regulatory battles have unfolded, and that is likely to remain so. The shaping of end-user behavior through rule and sanction was, and is, difficult. But intermediaries can be persuaded or required to shape users' technological experiences to channel them away from objectionable or illegal behavior, whether through hardware or operating system design of smart phones, or the shaping of software and services used by billions, such as by the most prominent social media platforms. The rights framework generally finds that such shaping should be limited, and in the late 1990s that was reflected in American law. For example, section 230 of the CDA – a part of the Act that remained after the Supreme Court struck down the rest – provided for immunity by platforms against many forms of potential liability occasioned by those platforms hosting and amplifying the speech of others, including end-users. And the notice-and-takedown safe harbors of the Digital Millennium Copyright Act offered a low-impact, routinized way for platforms to respond on a case-by-case basis to copyright complaints for others' material. Still, some scholars advocating for a rights framework thought these provisions went too far.

It was also in this rights-centric era that the Internet Corporation for Assigned Names and Numbers (ICANN) came about, chartered to bring consistency and "stakeholder" representation to policy-inflected decisions around global Internet naming and numbering, such as the number and nature of top-level domains (TLDs) like .com and .uk, including who would be charged with giving out or selling second-level names under those domains, and under what conditions. Apart from the simple desire to establish and regularize who would be earning money from the sale of domain names, the main concern aired as ICANN came into its own was about whether ICANN would itself become a censor of Internet content. ICANN could, the theory went, use its certification of TLD registries to, through a cascade of contracts, make for the suspension or transfer of domain names comprising or pointing to "bad stuff." Describing material in more precise terms of outright illegality has been difficult, since it would require a choice of which jurisdiction's definition of illegality to apply.

As it has happened, concerns about ICANN becoming the Internet police – infringing on individual rights – has so far seen ICANN's catalyzation of a suspension power to be only in the area of domain names whose very nature indicate a bad faith registration amounting to a form of trademark infringement. Domain names that are

not so infringing, but that are used as mnemonics for destinations containing harmful or illegal content, have generally not been touched by ICANN's policies.

The public health era

I was among those who celebrated the benefits of a rapidly-expanding Internet, both in scope and capability, thanks to the generative contributions of millions of users in code and content. For example, Internet protocols made possible the growth of the World Wide Web as an Internet application without any approvals sought or needed; the Web facilitated the rise of online wikis, and those wikis made possible the phenomenon of Wikipedia, which in turn invited contributions of content from people who themselves were not interested in coding software. Even amidst this celebration, in my case circa 2007, lay a new round of problems, which I described as part of the Generative Pattern:

1. An idea originates in a backwater.
2. It is ambitious but incomplete. It is partially implemented and released anyway, embracing the ethos of the procrastination principle.
3. Contribution is welcomed from all corners, resulting in an influx of usage.
4. Success is achieved beyond any expectation, and a higher profile draws even more usage.
5. Success is cut short: "There goes the neighborhood" as newer users are not conversant with the idea of experimentation and contribution, and other users are prepared to exploit the openness of the system to undesirable ends.
6. There is movement toward enclosure to prevent the problems that arise from the system's very popularity.

Indeed, the cutting short of success by those who subvert the system and take advantage of its now-many users – a problem arising from the very openness of the system itself – was recognized in earnest by 2010. (Though, as Whitney Phillips recently argued in a paper entitled "It Wasn't Just the Trolls: Early Internet Culture, 'Fun,' and the Fires of Exclusionary Laughter," the early days of internet culture were home to all manner of abuse, toxicity, and exclusion.) Cybersecurity had been my central worry; it was clear those problems were no longer wholesale, business-to-business issues, but something touching all of users' online activities. Without urgent attention given to developing a collective, generative defense, I worried about the Generative Pattern's conclusion: top-down enclosure to protect everyone by curtailing everyone's freedoms, demanded by the users themselves.

These kinds of concerns and how to meet them don't much benefit from a rights discourse, especially as they involve the mutual (if surely not symmetric) violation of rights by users against users, at least from a technical network point of view. Rather, they have much in common with how we talk about public health. They emphasize the interlinkages among us, the way that problems can all too easily spread from one person or node to another, and the need for systemic intervention and shaping to prevent harm from accruing, regardless of who might be to blame for first injecting harm into the system. Worries around viral malware hopping from one server to another have grown to be worries about mis- and disinformation hopping from one credulous person to another, abetted by social network intermediaries who

amplify controversial or outright false content if it increases user engagement with the platforms. Indeed, there is a literal public health dimension to misinformation today, as screeds and videos against even basic public vaccination, long proven to be beneficial, circulate and previously-near-defeated illnesses like measles make a startling comeback.

A public health framework is much more geared around risks and benefits than around individual rights. Pointing out harmful speech in a rights discourse might typically result in what amounts to a shrug and a declaration that such excesses are the "price of freedom," a sign that our commitment to rights requires sacrifice precisely where people would otherwise find the exercise of rights objectionable. In the public health frame, we instead are asked to gather empirical data about benefits and harms, and to brainstorm ways that the latter might be decreased without unduly trimming the former.

The process, or legitimacy, era

Reconciling rights and public health frameworks is not easy, not only between two people whose normative commitments fall into the respective camps, but also often within a single person: each framework can speak powerfully to us, favoring both individual liberty – including a skepticism over the responsible exercise of state power – while also sensitive to the fact that we live in a tightly-coupled, interlinked society, all the more so with the rise of networked technologies, and there are times when collective security calls for organized and perhaps even mindful architectural intervention. Moreover, the rise of intermediaries that not only facilitate communication with people we already know we want to reach – think email, or instant messaging – but also discovery of new ideas and people, means that there's a less-agreed-upon conception of neutrality or non-intervention. When Facebook or Twitter has millions of candidate items with which to salt a feed, any decision about what to show or recommend to you next is going to be freighted in a way that speeding delivery of a note between two discrete people is not.

We also happen to be in a time of very little trust in many if not most civic and private institutions, especially national and transnational ones. A simple vote in a legislature, or split decision from a court, seems not to well settle the complex and deeply debated issues that spring around digital governance.

This may be why we've lately seen some of today's most powerful private intermediaries, such as Facebook and Cloudflare, expressing uncertainty or contradiction about their own policies for intervention, a.k.a. intermeddling, vs. abstention, a.k.a. abdication. The rise of mainstream AI means that even detailed policies can be applied – or misapplied – in real time to the activities of billions of people so voluminous to otherwise be beyond moderation.

These companies have made some attempts to take decisions about content or user behavior out of their terms-of-service, customer support channels, and into some new institutional configuration meant to match the gravity of the questions around abuse, harassment, and the promotion or stifling of political speech. Facebook has proposed an independent oversight board, whose decisions would be binding

upon the company. Others have sought internal boards to reflect upon ethically-freighted decisions before making them. And regulators, loathe to try to make the decisions themselves at scale, have sought to require private intermediaries to impose particular standards without offering much by way of detail, such as in the current implementation of the European right to be forgotten.

What the field of digital governance, and indeed the world at large, needs, are ideas for new institutions and institutional relationships that can come to closure, however temporary, on some of these questions, and, like the project of law and political processes themselves, understand that all views will not and cannot be reconciled. But ideally even those who feel they have lost in a particular dispute or debate will not feel that they have been taken advantage of, or that the project to which they are contributing and are subject to – some digital expression of ideas and power – is not morally bankrupt.

The key to the next era of digital governance lies not in some abstract evaluation of whether our affordances are structured in ways that are correct or incorrect on one person's view, but rather if they are legitimate because of the inclusive and deliberative, and where possible, federated, way in which they were settled.

This contribution is based on Jonathan Zittrain, Three Eras of Digital Governance, in *Kleinwächter/Kettemann/Senges* (eds.), Towards a Global Framework for Cyber Peace and Digital Cooperation. An Agenda for the 2020s (Berlin: BMWi, November 2019), www.nextgenig.org.

Jonathan Zittrain *is the George Bemis Professor of International Law at Harvard Law School and the Harvard Kennedy School of Government, Professor of Computer Science at the Harvard School of Engineering and Applied Sciences, Director of the Harvard Law School Library, and Co-Founder of the Berkman Klein Center for Internet & Society.*

Cite as: Jonathan Zittrain, "Three eras of digital governance", *Völkerrechtsblog*, 27 November 2019.